Sequence-modification in copoly(ester-imide)s: a catalytic/supramolecular approach to the evolution and reading of copolymer sequence-information

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Sequence modification in copoly(ester-imide)s: a catalytic/supramolecular approach to the evolution and reading of copolymer sequence information

Marcus Knappert[1] · Howard M. Colquhoun [1]

Abstract

Catalytic ester-interchange reactions, analogous to mutation and recombination, allow new sequence information to be written statistically into poly(ester-imide) chains based on NDI (1,4,5,8-naphthalenetetracarboxylic diimide) units. Thus, both the insertion of the cyclic ester cyclopentadecanolide ("exaltolide") into an NDI-based homopolymer and quantitative sequence exchange between two different homopoly(ester-imide)s are catalyzed by di-n-butyl tin(IV) oxide. Emerging sequences are identified at the triplet and quintet levels using supramolecular complexation of pyrene-d_{10} at the NDI residues to amplify the separation of ^{1}H NMR resonances associated with different sequences. In such systems, pyrene is able to act as a "reader molecule" by generating different levels of ring-current shielding from the different patterns of supramolecular binding to all NDI-centered sequences of a given length.

Introduction

The processing of information in biological systems involves a vast and complex array of molecular machinery [1], most of which is devoted to "reading" copolymer sequence information [2], notably through the transcription (from DNA to RNA) [3] and translation (from RNA to protein) [4] of genetically relevant sections of polynucleotide chains. Any modification of existing DNA sequences (i.e., the "writing" of new DNA sequence information) is mainly carried out by statistical mechanisms such as mutation [5] and genetic recombination ("crossing over") [6]. Beneficial modifications of DNA are retained by natural selection at the organism level, and damaging modifications are discarded. Recombination involves a limited exchange of sequences between closely related segments of DNA chains originating in different parental genomes; mutation

can involve the random insertion or deletion of just a single nucleotide, but in chemical terms, both mutation and recombination are just interchange reactions at phosphodiester linkages in copolymer chains. Catalytic ester interchange is a well-established reaction in polymer chemistry [7–9]. Therefore, in the present work, we sought to develop synthetic analogs of mutation and recombination, complementing earlier work on "reading" comonomer sequences in copoly(ester-imide)s [10–13].

It has been reported that supramolecular complexation of a polycyclic aromatic molecule, such as pyrene, with diimide residues in binary copolyimides via complementary π-π-stacking [14], produces highly sequence-dependent NMR complexation shifts of the diimide resonances [12, 13]. This results from cumulative ring-current shielding [15] by the pyrene "reader molecules" as they bind at different sets of positions in different sequences and has allowed the resolution and assignment of diimide ^{1}H resonances to sequences containing up to seven diimide residues [12]. It has been shown in a previous publication [13] that such systems undergo fast exchange on the NMR timescale and that the resolution of the sequence information, as observed in the spectrum, is increased as the molar ratio of pyrene to NDI increases (see Supplementary Information). Increasing levels of pyrene drive the bound/unbound equilibrium toward the bound state, amplifying the ring-current shielding of NDI-based sequences by bound pyrene molecules.

✉ Howard M. Colquhoun
h.m.colquhoun@rdg.ac.uk

1 Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6DX, UK

Fig. 1 Monomers, homopolymers and copolymers described in the present work

Homopolymers 1 and 2 (x = 3 and 5)

Homopolymer 3

Copolymer 4

Copolymer 5

6   7   8
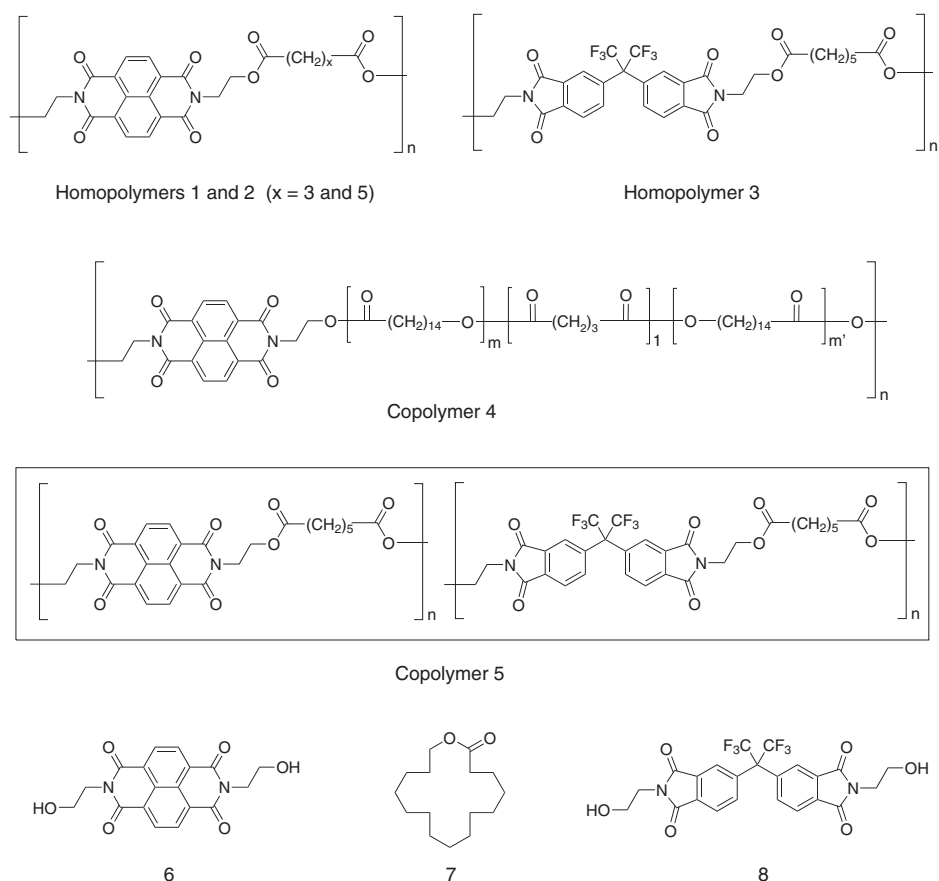
Sequence resolution at the five-residue level may be observed for copoly(*ester*-imide)s in which single-site supramolecular binding of pyrene to NDI has been demonstrated [13]. It is then possible that new sequence information might be "written" into such polymers by catalytic ester interchange, with the ring-current shielding method allowing new sequences to be "read" as they evolve. Here, we report the realization of this catalytic/supramolecular approach to the evolution of copolymer sequence information.

Results and discussion

Insertion chemistry

Homopoly(ester-imide) **1** (Fig. 1, M_n 57,000 g mol^{-1}), formed by high-temperature polycondensation of pentanedioyl (glutaryl) dichloride with diol **6** [12] based on the NDI (1,4,5,8-naphthalenetetracarboxylic diimide) unit, was found to react with cyclic ester exaltolide (**7**) at 180 °C in 1-chloronaphthalene as the solvent in the presence of di-*n*-butyl tin oxide as the catalyst. A 2:1 molar ratio for the repeat units between **1** and **7** was used, and the catalyst

concentration was 5 wt% on reactants. Analysis by ^{1}H NMR spectroscopy showed that after an hour, the single NDI resonance present in the spectrum of the starting homopolymer **1** had been replaced by a group of three closely spaced diimide resonances with estimated integrals of ca. 2:5:3 (high to low field, Fig. 2b). The addition of pyrene-d_{10} (5 equivalents per NDI residue) to the 1 h sample resulted in upfield shifts and increased separation of the three NDI resonances (Fig. 2c), allowing better definition of their relative integrals, although the presence of solvent and trace byproduct resonances limited the accuracy of integration. Samples of the product polymer at more extended reaction times showed no further change in the ^{1}H NMR spectrum, indicating that the insertion reaction was complete.

A theoretical analysis (see Supplementary Information, SI) of the potential for insertion of exaltolide (**7**) into homopolymer **1** showed that (i) reaction can occur at one or both ester linkages between the glutaric acid and diimide-diol residues, (ii) single or multiple exaltolide residues can be inserted at each of these positions, (iii) because of the symmetrical (diacid) nature of the glutaryl unit, the direction of exaltolide insertion must be reversed about the glutaryl residue (see Fig. 1), and (iv) for the same reason, only a single glutaryl residue can be present between any two
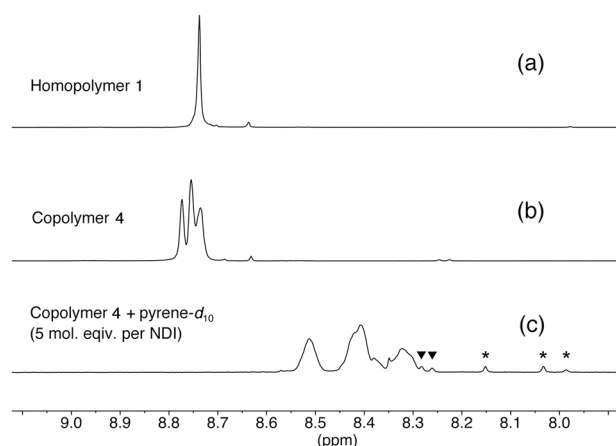
Fig. 2 ^{1}H NMR spectra for the NDI region of (**a**) homopolymer **1**; (**b**) copolymer **4**, formed by insertion of exaltolide (**7**) into homopolymer **1**; and (**c**) copolymer **4** in the presence of pyrene-d_{10} (5 mol equiv. per NDI residue). Starred resonances arise from residual protons in pyrene-d_{10}, and triangles indicate resonances from residual reaction solvent (1-chloronaphthalene). These and other impurity peaks limit the accuracy of integration in spectrum 1c, but relative values for the three principal resonances in spectrum 1c are estimated to be 2.0: 5.1: 2.4 (high to low field)

Table 1 Calculation of sequence probabilities in copolymer **4**

Sequence	Residue probabilities	Sequence probability
IGIGI	$1 \times (0.667)^4$	0.198
I(E$_n$GE$_{n'}$)IGI	$1 \times (0.833) \times (0.667)^3$	0.247
IGI(E$_n$GE$_{n'}$)I	$1 \times (0.833) \times (0.667)^3$	0.247
I(E$_n$GE$_{n'}$)I(E$_n$GE$_{n'}$)I	$1 \times (0.833)^2 \times (0.667)^2$	0.308
Total probability		1.000

successive NDI residues in the final copolymer (**4**). The structure of **4** can, thus, be abbreviated as [-I(E$_m$)G(E$_{m'}$)-]$_n$, where I = NDI-diol, E = exaltolide, and G = glutaryl residues. The numbers of exaltolide repeat units m and m' in **4** can take any value, including zero. If both are zero, then we simply have the starting material, [-I-G-]$_n$, homopolymer **1**. Consequently, the formulation (E$_m$GE$_{m'}$) describes *all possible* linking groups between any two "I" residues in both homopolymer **1** and copolymer **4**.

The simplest interpretation of the three NDI resonances in the ^{1}H NMR spectrum of copolymer **4** is that they represent different copolymer sequences derived from the starting "three-I" sequence [-IGIGI-]. As discussed in earlier papers [11–13], only the central diimide resonance needs to be considered in this type of analysis because the outer diimide residues are located at the centers of the other, analogous sequences and so can be treated separately. After insertion of exaltolide, a "central" diimide residue in copolymer **4** may, thus, be contained in four sequences: [-IGIGI-], [-I(E$_m$GE$_{m'}$)IGI-], [-IGI(E$_m$GE$_{m'}$)I-] and [-I(E$_m$GE$_{m'}$)I(E$_m$GE$_{m'}$)I-]. The

second and third sequences are directionally degenerate, and, thus, indistinguishable in the ^{1}H NMR spectrum, leaving just three potential NDI resonances for these four sequences. This analysis is valid when the copolymer molecular weight is high enough for end-group effects to be ignored, and, indeed, the integration of NDI-end-group resonances in the ^{1}H NMR spectrum of copolymer **4** against analogous in-chain resonances gives a M_n value of 27,000 g mol^{-1}, corresponding to a degree of polymerization of ca. 45, and, thus, to <5 mol% of end groups.

The general linking group (E$_m$GE$_{m'}$) between successive NDI residues can of course represent many different chemical entities, depending on the values of m and m', but as "E" is a linear, sixteen-atom moiety, even a single exaltolide insertion can more than double the number of atoms between successive I residues. Consequently, as far as the central NDI residue is concerned, the only significant distinguishing feature between its different environments is whether there are zero, one, or two "E" residues present in the units linking it to its neighboring NDIs. Thus, it only remains to calculate the probabilities of the sequences [-IGIGI-], [-I(E$_m$GE$_{m'}$)IGI-]/[-IGI(E$_m$GE$_{m'}$)I-] and [-I(E$_m$GE$_{m'}$)I(E$_m$GE$_{m'}$)I-] to predict the relative intensities (i.e., integrals) of the corresponding ^{1}H NMR resonances.

It can be shown (see Supplementary Information) that, in copolymer **4**, the probability of a linker between two **I**'s being simply G (i.e., m = m' = 0) is 0.667, and the probability of the linker being (E$_m$GE$_{m'}$) for all other values of m or m', taken together, is 0.833. The probability of an I, rather than an E being found at the end of each sequence, is also 0.667, but the probability of the central I must, by definition (as we are considering only I-centered sequences), be 1. Calculations of the individual sequence probabilities for copolymer **4** are given in Table 1.

Since the total probability, as shown, equals 1, this analysis must include all the probabilities for all the sequences in copolymer **4** that are based on three successive NDI residues. The relative intensities of NDI resonances in copolymer **4** are, thus, predicted to be 0.198:0.494:0.308. These values are reasonably close to those found experimentally in the ^{1}H NMR spectrum of the insertion product, copolymer **4**, confirming that an effective sequence-modification reaction analogous to mutation was achieved, and providing assignments for the three observed resonances to the sequences [-IGIGI-], [-I(E$_m$GE$_{m'}$)IGI-]/[-IGI(E$_m$GE$_{m'}$)I-] and [-I(E$_m$GE$_{m'}$)I(E$_m$GE$_{m'}$)I-] (high to low field, respectively).
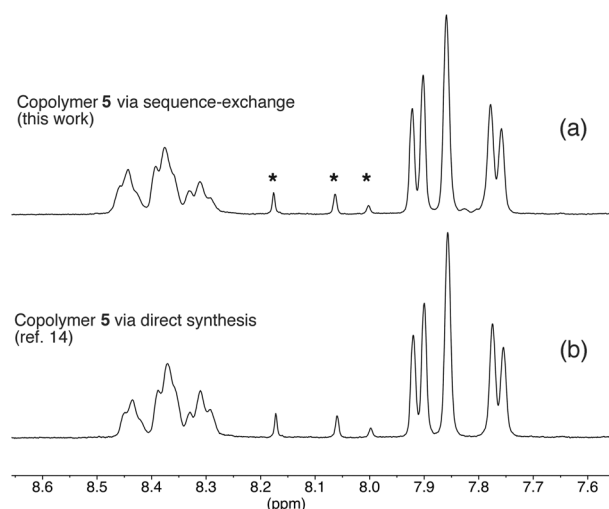
Sequence-exchange chemistry

Equimolar proportions (in diimide residues) of homopoly (ester-imide)s **2** and **3** (Fig. 1, M_n values 21,000 and 20,000 g mol^{-1}, respectively) were heated at 180 °C under
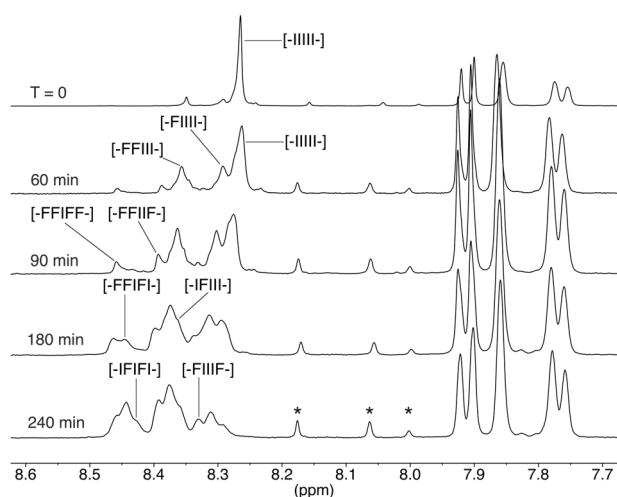
Fig. 3 ^{1}H NMR spectra showing the progressive development of new quintet sequences resulting from catalytic interchange of sequences between homopolymers **2** and **3**—a process ultimately leading (240 min) to the formation of copolymer **5**. "I" = NDI and "F" = HFDI. Peaks between 7.70 ppm and 7.95 ppm are due to HFDI protons, and the starred resonances between 8.00 ppm and 8.20 ppm represent residual protons in pyrene-d_{10} (7 equivalents added per mole of NDI residues). For peak assignment methodology, see ref. 13, Table 1)

Fig. 4 ^{1}H NMR spectra for copolymer **5** in the presence of pyrene-d_{10}. (**a**) Sample from ester interchange (this work) and (**b**) sample from direct synthesis ($M_n = 22,000$) [13]. Peaks between 7.70 ppm and 7.95 ppm are due to HFDI ("F") protons, and the starred resonances between 8.00 ppm and 8.20 ppm represent residual protons in (99.8% D) pyrene-d_{10}. Peaks between 8.2 ppm and 8.5 ppm are due to NDI ("I") protons centered in a range of different sequences. In both spectra, the nine resolved resonances are assigned (highest to lowest field) to I-centered quintet sequences: IIIII; FIIII/IIIIF; FIIIF; IFIII/IIIFI; FFIII/IIIFF/FIIFI/IFIIF; FFIIF/FIIFF; IFIFI; FFIFI/IFIFF; FFIFF [13]

nitrogen in 1,2-dichlorobenzene as the solvent in the presence of di-*n*-butyl tin(IV) oxide as the catalyst. Samples of the reaction mixture were taken at 10 min intervals, and the polymeric component was isolated and analyzed by ^{1}H NMR spectroscopy. Initially, the ^{1}H NMR spectrum of a mixture of the two starting homopolymers showed just a single NDI resonance: the addition of pyrene-d_{10} (7 equivalents per NDI) to the NMR sample shifted this signal upfield by ca. 0.5 ppm, but no splitting of the resonance was observed (Fig. 3). However, after a reaction time of one hour, new resonances appeared at a lower field (Fig. 3; again using pyrene addition to expand the spectrum), and from previous work, these resonances were assigned to sequences containing both NDI and HFDI residues [14]. After 2 h, further new resonances were evident at still lower field, and the highest-field (all-NDI) resonance diminished substantially in intensity. Finally, after 4 h, the original (single) NDI peak was replaced by no fewer than nine NDI resonances in the range 8.25–8.50 ppm (spectrum measured in the presence of seven equivalents of pyrene-d_{10} per NDI), with approximate relative intensities of 1:2:1:2:4:2:1:2:1 (Fig. 3). This final pattern was previously identified as being characteristic of random, binary copolymer **5**, synthesized by statistical copolycondensation of monomers **6** and **8** (equimolar ratio) with heptanedioyl (pimeloyl) dichloride [13].

An earlier analysis of this nine-line pattern [13], based on single-site binding of pyrene to each NDI ("I") residue in the copolymer, showed that it arises from a fractal distribution of ring-current shieldings that sum to different

values for the central diimide residue in one or more of the different sequences present. This earlier analysis [13] allowed all nine resonances to be assigned to specific NDI-centered quintets (or groups of quintets), so that here, as shown in Fig. 3, we can follow in detail the emergence of new sequences resulting from ester interchange between the starting homopolymers **2** and **3**.

Finally, the ^{1}H NMR spectra for samples of copolymer **5** obtained by (a) direct copolymerisation [13] and (b) sequence exchange are compared in Fig. 4, using pyrene-d_{10} to resolve sequence information at the quintet level. The same sequences are clearly represented in the two spectra, even if there are slight differences between the two copolymers in terms of their relative sequence populations. It is evident from this result that catalytic sequence exchange can indeed provide a successful approach to the statistical "writing" of new sequences in poly(ester-imide) systems.

Conclusions

Catalytic ester-interchange reactions, analogous to mutation and recombination, enable new sequence information to be written statistically into NDI-based poly(ester-imide) chains. Emerging sequences may be identified at the triplet and quintet levels by ^{1}H NMR analysis using supramolecular complexation of pyrene-d_{10} to the diimide residues. In such systems, pyrene acts as a "reader molecule" by

generating different levels of ring-current shielding from the different patterns of binding to all possible NDI sequences of any given length.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Goodsell DS. The Machinery of Life. 2nd ed. New York: Springer; 2009.
2. Adami C. Information theory in molecular biology. Phys Life Rev. 2004;1:3–22.
3. Bushnell DA, Westover KD, Davis RE, Kornberg RD. Structural basis of transcription: an RNA polymerase II–TFIIB cocrystal at 4.5 angstroms. Science. 2004;303:983–8.
4. Gualerzi CO, Pon CL. Initiation of mRNA translation in prokaryotes. Biochemistry. 1990;29:5881–9.
5. Crick FHC, Barnett L, Brenner S, Watts-Tobin RJ. General nature of the genetic code for proteins. Nature. 1961;192:1227–32.
6. Holliday R. Recombination and meiosis. Philos Trans R Soc B. 1977;277:359–70.
7. Shuster M, Narkis M, Siegmann A. Catalytic transesterification in polycarbonate- polycaprolactone systems. J Appl Polym Sci. 1994;52:1383–91.
8. Hodge P, Yang Z, Ben-Haida A, McGrail CS. Cyclodepolymerisation of poly (ethylene naphthalene-2,6-dicarboxylate) and ring-opening polymerisations of the cyclic oligomers obtained. J Mater Chem. 2000;10:1533–7.
9. González-Vidal N, Martínez de Ilarduya A, Herrera V, Muñoz-Guerra S. Poly(hexamethylene terephthalate-co-caprolactone) copolyesters obtained by ring- opening polymerization. Macromolecules. 2008;41:4136–46.
10. Zhu Z, Cardin CJ, Gan Y, Colquhoun HM. Sequence-selective assembly of tweezer- molecules on linear templates enables frameshift reading of sequence information. Nat Chem. 2010;2:653–60.
11. Shaw JS, Vaiyapuri R, Parker MP, Murray CA, Lim KJC, Pan C, et al. Elements of fractal geometry in the ^{1}H NMR spectrum of a copolymer intercalation-complex: identification of the underlying Cantor set. Chem Sci. 2018;9:4052–61.
12. Knappert M, Jin T, Midgley SD, Wu G, Scherman OA, Grau-Crespo R, et al. Supramolecular complexation between chain-folding poly(ester-imide)s and polycyclic aromatics: a fractal-based pattern of NMR ring-current shielding. Polym Chem. 2019;10:6641–50.
13. Knappert M, Jin T, Midgley SD, Wu G, Scherman OA, Grau-Crespo R, et al. Single-site binding of pyrene to poly(ester-imide)s incorporating long spacer-units: Prediction of NMR resonance-patterns from a fractal model. Chem Sci. 2020;11:12165–77.
14. Hunter CA, Sanders JKM. The nature of π-π interactions. J Am Chem Soc. 1990;112:5525–34.
15. Klod S, Kleinpeter E. Ab initio calculation of the anisotropy effect of multiple bonds and the ring current effect of arenes – application in conformational and configurational analysis. J Chem Soc, Perkin Trans. 2001;2:1893–8.